

1633

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/202,047

DATE: 07/21/2000
TIME: 11:56:47

Input Set : A:\0020-4491P.APP
Output Set: N:\CRF3\07212000\I202047.raw

ENTERED

RECEIVED

JUL 26 2000

TECH CENTER 1600/2900

```

3 <110> APPLICANT: ITOH, Kyogo
4   SHICHIJO, Shigeki
5   IMAI, Yasuhisa
7 <120> TITLE OF INVENTION: TUMOR ANTIGEN PROTEINS, GENES THEREFOR, AND TUMOR
8   ANTIGEN PEPTIDES
10 <130> FILE REFERENCE: 0020-4491P
12 <140> CURRENT APPLICATION NUMBER: 09/202,047
13 <141> CURRENT FILING DATE: 1998-12-07
15 <160> NUMBER OF SEQ ID NOS: 2
17 <170> SOFTWARE: PatentIn Ver. 2.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 2527
21 <212> TYPE: DNA
22 <213> ORGANISM: Homo sapiens
24 <220> FEATURE:
25 <221> NAME/KEY: 5'UTR
26 <222> LOCATION: (1)..(38)
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (39)..(2438)
32 <220> FEATURE:
33 <221> NAME/KEY: 3'UTR
34 <222> LOCATION: (2439)..(2506)
36 <220> FEATURE:
37 <221> NAME/KEY: polyA_site
38 <222> LOCATION: (2507)..(2527)
40 <400> SEQUENCE: 1
41 gggtcggcg cagccgggct cggagtggtgac gtgccaact atg ggg tcg tcc aag aag 56
42                                     Met Gly Ser Ser Lys Lys
43                                     1           5
45 cat cgc gga gag aag gag gcg gcc ggg acg acg gcg gcg gcc ggc acc 104
46 His Arg Gly Glu Lys Glu Ala Ala Gly Thr Thr Ala Ala Ala Gly Thr
47           10           15           20
49 ggg ggt gcc acc gag cag ccg ccg cgg cac cgg gaa cac aaa aaa cac 152
50 Gly Gly Ala Thr Glu Gln Pro Pro Arg His Arg Glu His Lys Lys His
51           25           30           35
53 aag cac cgg agt ggc ggc agt ggc ggt agc ggt ggc gaa cga cgg aag 200
54 Lys His Arg Ser Gly Gly Ser Gly Gly Ser Gly Gly Glu Arg Arg Lys
55           40           45           50
57 cgg agc cgg gaa cgt ggg ggc gag cgc ggg agc ggg cgg cgc ggg gcc 248
58 Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly Ser Gly Arg Arg Gly Ala
59 55           60           65           70
61 gaa gct gag gcc cgg agc agc acg cac ggg cgg gag cgc agc cag gca 296
62 Glu Ala Glu Ala Arg Ser Ser Thr His Gly Arg Glu Arg Ser Gln Ala
63           75           80           85
65 gag ccc tcc gag cgg cgc gtg aag cgg gag aag cgc gat gac gcc tac 344
66 Glu Pro Ser Glu Arg Arg Val Lys Arg Glu Lys Arg Asp Asp Gly Tyr

```

RAW SEQUENCE LISTING

DATE: 07/21/2000

PATENT APPLICATION: US/09/202,047

TIME: 11:56:47

Input Set : A:\0020-4491P.APP

Output Set: N:\CRF3\07212000\I202047.raw

```

67          90          95          100
69 gag gcc gct gcc agc tcc aaa act agc tca ggc gat gcc tcc tca ctc 392
70 Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser Gly Asp Ala Ser Ser Leu
71          105          110          115
73 agc atc gag gag act aac aaa ctc cgg gca aag ttg ggg ctg aaa ccc 440
74 Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala Lys Leu Gly Leu Lys Pro
75          120          125          130
77 ttg gag gtt aat gcc atc aag aag gag gcg ggc acc aag gag gag ccc 488
78 Leu Glu Val Asn Ala Ile Lys Lys Glu Ala Gly Thr Lys Glu Glu Pro
79 135          140          145          150
81 gtg aca gct gat gtc atc aac cct atg gcc ttg cga cag cga gag gag 536
82 Val Thr Ala Asp Val Ile Asn Pro Met Ala Leu Arg Gln Arg Glu Glu
83          155          160          165
85 ctg cgg gag aag ctg gcg gct gcc aag gag aag cgc ctg ctg aac caa 584
86 Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu Lys Arg Leu Leu Asn Gln
87          170          175          180
89 aag ctg ggg aag ata aag acc cta gga gag gat gac ccc tgg ctg gac 632
90 Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu Asp Asp Pro Trp Leu Asp
91          185          190          195
93 gac act gca gcc tgg atc gag agg gcc agc cgg cag ctg cag aag gag aag 680
94 Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg Gln Leu Gln Lys Glu Lys
95          200          205          210
97 gac ctg gca gag aag agg gcc aag tta ctg gag gag atg gac caa gag 728
98 Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu Glu Glu Met Asp Gln Glu
99 215          220          225          230
101 ttt ggt gtc agc act ctg gtg gag gag gag ttc ggg cag agg cgg cag 776
102 Phe Gly Val Ser Thr Leu Val Glu Glu Glu Phe Gly Gln Arg Arg Gln
103          235          240          245
105 gac ctg tac agt gcc cgg gac ctg cag ggc ctc acc gtg gag cat gcc 824
106 Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly Leu Thr Val Glu His Ala
107          250          255          260
109 att gat tcc ttc cga gaa ggg gag aca atg att ctt acc ctc aag gac 872
110 Ile Asp Ser Phe Arg Glu Gly Glu Thr Met Ile Leu Thr Leu Lys Asp
111          265          270          275
113 aaa ggc gtg ctg cag gag gag gag gac gtg ctg gtg aac gtg aac ctg 920
114 Lys Gly Val Leu Gln Glu Glu Glu Asp Val Leu Val Asn Val Asn Leu
115          280          285          290
117 gtg gat aag gag cgg gca gag aaa aat gtg gag ctg cgg aag aag aag 968
118 Val Asp Lys Glu Arg Ala Glu Lys Asn Val Glu Leu Arg Lys Lys Lys
119 295          300          305          310
121 cct gac tac ctg ccc tat gcc gag gac gag agc gtg gac gac ctg gcg 1016
122 Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu Ser Val Asp Asp Leu Ala
123          315          320          325
125 cag caa aaa cct cgc tct atc ctg tcc aag tat gac gaa gag ctt gaa 1064
126 Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys Tyr Asp Glu Glu Leu Glu
127          330          335          340
129 ggg gag cgg cca cat tcc ttc cgc ttg gag cag ggc ggc acg gct gat 1112
130 Gly Glu Arg Pro His Ser Phe Arg Leu Glu Gln Gly Gly Thr Ala Asp
131          345          350          355

```

RAW SEQUENCE LISTING

DATE: 07/21/2000

PATENT APPLICATION: US/09/202,047

TIME: 11:56:47

Input Set : A:\0020-4491P.APP

Output Set: N:\CRF3\07212000\I202047.raw

```

133 ggc ctg cgg gag cgg gag ctg gag gag atc cgg gcc aag ctg cgg ctg 1160
134 Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile Arg Ala Lys Leu Arg Leu
135 360 365 370
137 cag gct cag tcc ctg agc aca gtg ggg ccc cgg ctg gcc tcc gaa tac 1208
138 Gln Ala Gln Ser Leu Ser Thr Val Gly Pro Arg Leu Ala Ser Glu Tyr
139 375 380 385 390
141 ctc acg cct gag gag atg gtg acc ttt aaa aag acc aag cgg agg gtg 1256
142 Leu Thr Pro Glu Glu Met Val Thr Phe Lys Lys Thr Lys Arg Arg Val
143 395 400 405
145 aag aaa atc cgc aag aag gag aag gag gta gta gtg cgg gca gat gac 1304
146 Lys Lys Ile Arg Lys Lys Glu Lys Glu Val Val Val Arg Ala Asp Asp
147 410 415 420
149 ttg ctg cct ctc ggg gac cag act cag gat ggg gac ttt ggt tcc aga 1352
150 Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp Gly Asp Phe Gly Ser Arg
151 425 430 435
153 ctg cgg gga cgg ggt cgc cgc cga gtg tcc gaa gtg gag gag gag aag 1400
154 Leu Arg Gly Arg Gly Arg Arg Val Ser Glu Val Glu Glu Glu Lys
155 440 445 450
157 gag cct gtg cct cag ccc ctg ccg tcg gac gac acc cga gtg gag aac 1448
158 Glu Pro Val Pro Gln Pro Leu Pro Ser Asp Thr Arg Val Glu Asn
159 455 460 465 470
161 atg gac atc agt gat gag gag gaa ggt gga gct cca ccg ccg ggg tcc 1496
162 Met Asp Ile Ser Asp Glu Glu Glu Gly Gly Ala Pro Pro Pro Gly Ser
163 475 480 485
165 ccg cag gtg ctg gag gag gac gag gcg gag ctg gag ctg cag aag cag 1544
166 Pro Gln Val Leu Glu Glu Asp Glu Ala Glu Leu Glu Leu Lys Gln
167 490 495 500
169 ctg gag aag gga cgc cgg ctg cga cag tta cag cag cta cag cag ctg 1592
170 Leu Glu Lys Gly Arg Arg Leu Arg Gln Leu Gln Gln Leu Gln Gln Leu
171 505 510 515
173 cga gac agt ggc gag aag gtg gtg gag att gtg aag aag ctg gag tct 1640
174 Arg Asp Ser Gly Glu Lys Val Val Glu Ile Val Lys Lys Leu Glu Ser
175 520 525 530
177 cgc cag cgg ggc tgg gag gag gat gag gat ccc gag cgg aag ggg gcc 1688
178 Arg Gln Arg Gly Trp Glu Glu Asp Glu Asp Pro Glu Arg Lys Gly Ala
179 535 540 545 550
181 atc gtg ttc aac gcc acg tcc gag ttc tgc cgc acc ttg ggg gag atc 1736
182 Ile Val Phe Asn Ala Thr Ser Glu Phe Cys Arg Thr Leu Gly Glu Ile
183 555 560 565
185 ccc acc tac ggg ctg gct ggc aat cgc gag gag cag gag gag ctc atg 1784
186 Pro Thr Tyr Gly Leu Ala Gly Asn Arg Glu Glu Gln Glu Glu Leu Met
187 570 575 580
189 gac ttt gaa cgg gat gag gag cgc tca gcc aac ggt ggc tcc gaa tct 1832
190 Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala Asn Gly Gly Ser Glu Ser
191 585 590 595
193 gac ggg gag gag aac atc ggc tgg agc acg gtg aac ctg gac gag gag 1880
194 Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr Val Asn Leu Asp Glu Glu
195 600 605 610
197 aag cag cag cag gat ttc tct gct tcc tcc acc acc atc ctg gac gag 1928

```

RECEIVED
TECH CENTER 1600/2930

RAW SEQUENCE LISTING

DATE: 07/21/2000

PATENT APPLICATION: US/09/202,047

TIME: 11:56:47

Input Set : A:\0020-4491P.APP

Output Set: N:\CRF3\07212000\I202047.raw

```

198 Lys Gln Gln Gln Asp Phe Ser Ala Ser Ser Thr Thr Ile Leu Asp Glu
199 615 620 625 630
201 gaa ccg atc gtg aat agg ggg ctg gca gct gcc ctg ctc ctg tgt cag 1976
202 Glu Pro Ile Val Asn Arg Gly Leu Ala Ala Leu Leu Leu Cys Gln
203 635 640 645
205 aac aaa ggg ctg ctg gag acc aca gtg cag aag gtg gcc cgg gtg aag 2024
206 Asn Lys Gly Leu Leu Glu Thr Thr Val Gln Lys Val Ala Arg Val Lys
207 650 655 660
209 gcc ccc aac aag tcg ctg ccc tca gcc gtg tac tgc atc gag gat aag 2072
210 Ala Pro Asn Lys Ser Leu Pro Ser Ala Val Tyr Cys Ile Glu Asp Lys
211 665 670 675
213 atg gcc atc gat gac aag tac agc cgg agg gag gaa tac cga ggc ttc 2120
214 Met Ala Ile Asp Asp Lys Tyr Ser Arg Arg Glu Glu Tyr Arg Gly Phe
215 680 685 690
217 aca cag gac ttc aag gag aag gac ggc tac aaa ccc gac gtt aag atc 2168
218 Thr Gln Asp Phe Lys Glu Lys Asp Gly Tyr Lys Pro Asp Val Lys Ile
219 695 700 705 710
221 gaa tac gtg gat gag acg ggc cgg aaa ctc aca ccc aag gag gct ttc 2216
222 Glu Tyr Val Asp Glu Thr Gly Arg Lys Leu Thr Pro Lys Glu Ala Phe
223 715 720 725
225 cgg cag ctg tcg cac cgc ttc cat ggc aag ggc tca ggc aag atg aag 2264
226 Arg Gln Leu Ser His Arg Phe His Gly Lys Gly Ser Gly Lys Met Lys
227 730 735 740
229 aca gag cgg cgg atg aag aag ctg gac gag gag gcg ctc ctg aag aag 2312
230 Thr Glu Arg Arg Met Lys Lys Leu Asp Glu Glu Ala Leu Leu Lys Lys
231 745 750 755
233 atg agc tcc agc gac acg ccc ctg ggc acc gtg gcc ctg ctc cag gag 2360
234 Met Ser Ser Ser Asp Thr Pro Leu Gly Thr Val Ala Leu Leu Gln Glu
235 760 765 770
237 aag cag aag gct cag aag acc ccc tac atc gtg ctc agc ggc agc ggc 2408
238 Lys Gln Lys Ala Gln Lys Thr Pro Tyr Ile Val Leu Ser Gly Ser Gly
239 775 780 785 790
241 aag agc atg aac gcg aac acc atc acc aag tgacagcgcc ctcccgtagt 2458
242 Lys Ser Met Asn Ala Asn Thr Ile Thr Lys
243 795 800
245 cggcctgcc tcaaccttca tattaataaa agctccctcc ttatttttaa aaaaaaaaaa 2518
247 aaaaaaaaaa 2527
250 <210> SEQ ID NO: 2
251 <211> LENGTH: 800
252 <212> TYPE: PRT
253 <213> ORGANISM: Homo sapiens
255 <400> SEQUENCE: 2
256 Met Gly Ser Ser Lys Lys His Arg Gly Glu Lys Glu Ala Ala Gly Thr
257 1 5 10 15
259 Thr Ala Ala Ala Gly Thr Gly Gly Ala Thr Glu Gln Pro Pro Arg His
260 20 25 30
262 Arg Glu His Lys Lys His Lys His Arg Ser Gly Gly Ser Gly Gly Ser
263 35 40 45
265 Gly Gly Glu Arg Arg Lys Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/202,047

DATE: 07/21/2000

TIME: 11:56:47

Input Set : A:\0020-4491P.APP

Output Set: N:\CRF3\07212000\I202047.raw

```

266      50      55      60
268 Ser Gly Arg Arg Gly Ala Glu Ala Glu Ala Arg Ser Ser Thr His Gly
269 65      70      75      80
271 Arg Glu Arg Ser Gln Ala Glu Pro Ser Glu Arg Arg Val Lys Arg Glu
272      85      90      95
274 Lys Arg Asp Asp Gly Tyr Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser
275      100      105      110
277 Gly Asp Ala Ser Ser Leu Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala
278      115      120      125
280 Lys Leu Gly Leu Lys Pro Leu Glu Val Asn Ala Ile Lys Lys Glu Ala
281      130      135      140
283 Gly Thr Lys Glu Glu Pro Val Thr Ala Asp Val Ile Asn Pro Met Ala
284 145      150      155      160
286 Leu Arg Gln Arg Glu Glu Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu
287      165      170      175
289 Lys Arg Leu Leu Asn Gln Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu
290      180      185      190
292 Asp Asp Pro Trp Leu Asp Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg
293      195      200      205
295 Gln Leu Gln Lys Glu Lys Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu
296      210      215      220
298 Glu Glu Met Asp Gln Glu Phe Gly Val Ser Thr Leu Val Glu Glu Glu
299 225      230      235      240
301 Phe Gly Gln Arg Arg Gln Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly
302      245      250      255
304 Leu Thr Val Glu His Ala Ile Asp Ser Phe Arg Glu Gly Glu Thr Met
305      260      265      270
307 Ile Leu Thr Leu Lys Asp Lys Gly Val Leu Gln Glu Glu Glu Asp Val
308      275      280      285
310 Leu Val Asn Val Asn Leu Val Asp Lys Glu Arg Ala Glu Lys Asn Val
311      290      295      300
313 Glu Leu Arg Lys Lys Lys Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu
314 305      310      315      320
316 Ser Val Asp Asp Leu Ala Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys
317      325      330      335
319 Tyr Asp Glu Glu Leu Glu Gly Glu Arg Pro His Ser Phe Arg Leu Glu
320      340      345      350
322 Gln Gly Gly Thr Ala Asp Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile
323      355      360      365
325 Arg Ala Lys Leu Arg Leu Gln Ala Gln Ser Leu Ser Thr Val Gly Pro
326      370      375      380
328 Arg Leu Ala Ser Glu Tyr Leu Thr Pro Glu Glu Met Val Thr Phe Lys
329 385      390      395      400
331 Lys Thr Lys Arg Arg Val Lys Lys Ile Arg Lys Lys Glu Lys Glu Val
332      405      410      415
334 Val Val Arg Ala Asp Asp Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp
335      420      425      430
337 Gly Asp Phe Gly Ser Arg Leu Arg Gly Arg Gly Arg Arg Val Ser
338      435      440      445

```

VERIFICATION SUMMARY DATE: 07/21/2000
PATENT APPLICATION: US/09/202,047 TIME: 11:56:48

Input Set : A:\0020-4491P.APP
Output Set: N:\CRF3\07212000\I202047.raw